Influence Diagnostic for Semiparametric Partially Nonlinear Mixed-Effects Models

Robson Jose Mariano Machado*
University College London, London, United Kingdom - robson.machado.14@ucl.ac.uk

Cibele Maria Russo
Universidade de São Paulo, São Carlos, Brazil - cibele@icmc.usp.br

Abstract

The aim of this paper is to develop local influence analysis in a semiparametric partially nonlinear mixed-effects models. This model generalises the parametric nonlinear mixed-effects models by including a nonparametric function to smooth the mean response curve. Furthermore, the random effects are included linearly to the model, which provides computational advantages to the estimating procedure, as estimation is usually complicated in nonlinear mixed-effects models. The proposed methodology is illustrated with a pharmacokinetic data and the local influence analysis is performed to identify potential influential observations.

Keywords: Nonlinear models; Mixed-effects models; Local influence; Smoothing.

1. Introduction

Semiparametric partially nonlinear mixed-effects models are proposed for analysing correlated data with a nonlinear structure. This type of models assume that the respond variable is explained by a nonlinear function of fixed effects parameters and explanatory variables, a nonparametric function that provides flexibility to the functional form that underlies the data and random errors. In the proposed model, the random effects are included linearly to the model so that it is straightforward to work with the marginal model. In this paper, the proposed model is applied to the theophylline data, previously analysed by Machado and Russo (2013), in which the drug was administered to 12 subjects and serum concentrations were measured at 11 time points over the 25 hours after being administered. Finally, a local influence analysis is developed to identify potential influential observations.

2. The model

A semiparametric partially nonlinear mixed-effects model for the jth response of the ith subject \( y_{ij} \) may be written as

\[
y_{ij} = \eta(x_{ij}, \beta) + f(t_{ij}) + \mathbf{Z}_{ij}^\top \mathbf{u}_i + \epsilon_{ij}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, m_i,
\]

where \( \eta \) is a nonlinear function of \( \beta \) and \( x_{ij} \), \( x_{ij} \) is a covariate that can be a scalar or a vector, \( \mathbf{Z}_{ij} \) is a known vector related to the random effects, \( \beta = (\beta_1, \ldots, \beta_p) \) a vector of unknown parameters of major interest, \( \mathbf{u}_i = (u_{i1}, \ldots, u_{ip}) \) the random-effects coefficients for the ith subject, \( f \) is a smooth function of time, \( t_{ij} \) are the time points. In this work the smooth function is a \( B \)-spline of degree 3 with \( L \) equidistant knots. In matrix notation, the model (1) for a vector of longitudinal response variable \( \mathbf{y}_i(m_i \times 1) \) may be expressed as follows

\[
\mathbf{y}_i = \eta(\mathbf{x}_i, \beta) + \mathbf{B}_i \mathbf{u}_i + \mathbf{Z}_i \mathbf{u}_i + \epsilon_i, \quad i = 1, \ldots, n,
\]

where \( \mathbf{y}_i = (y_{i1}, \ldots, y_{im_i})^\top \), \( \mathbf{x}_i = (x_{i1}, \ldots, x_{im_i})^\top \), \( \mathbf{Z}_i = (\mathbf{Z}_{i1}, \ldots, \mathbf{Z}_{im_i})^\top \), \( \eta(\mathbf{x}_i, \beta) = (\eta(x_{i1}, \beta), \ldots, \eta(x_{im_i}, \beta))^\top \), the elements of matrix \( \mathbf{B}_i \) are \( b_{ij} = B_k(t_{ij}) \) for the ith subject and \( \mathbf{u}_i = (\alpha_1, \ldots, \alpha_L)^\top \) is vector of \( B \)-splines coefficients. It is assumed that \( \mathbf{u}_i \) and \( \epsilon_i \) are independent and follow a multivariate normal distribution. It is easy in this case to work on the marginal model, \( \mathbf{y}_i \sim N_{m_i}(\eta(\mathbf{x}_i, \beta) + \mathbf{B}_i \mathbf{u}_i; \Sigma_i) \), to preserve the mean of the hierarchical model without requiring numerical integration. The penalised log-likelihood function based on \( P \)-splines is given by

\[
L_p(\beta, \alpha, \Sigma_i; \mathbf{y}_i, \mathbf{x}_i) = L(\beta, \alpha, \Sigma_i; \mathbf{y}_i, \mathbf{x}_i) - \frac{\lambda}{2n} \alpha^\top \mathbf{D}_k^\top \mathbf{D}_k \alpha,
\]

where...
in which \( L_i(\beta, \alpha, \Sigma; y, x) \) is the log-likelihood function to the \( i \)th subject, \( \lambda \) is a positive real number to be estimated, \( D_i \) is the matrix representation of the difference operator and \( k \) is the order of the differences. The complete penalised log-likelihood is given by \( L_p(\beta, \alpha, \Sigma; y, x) = \sum_{i=1}^{n} L_p. \) Therefore, it is possible to estimate the model by the Fisher scoring algorithm (Poon & Poon 1996).

### 3. Local influence

Consider the smoothing parameter fixed. Let \( L_p(\theta, \omega) \) be the log-likelihood function to the model, in which \( \omega \) is a vector of parameters. The perturbations are introduced to the model through the vector \( \omega \) that belongs to the open set \( \Omega \in \mathbb{R}^k. \) Let \( L_p(\theta, \omega) \) the penalised log-likelihood function corresponding to perturbed model. It is assumed that there exists a \( \omega_0 \in \Theta \) such as \( L_p(\theta, \omega) = L_p(\theta, \omega_0). \) To assess the local influence of a perturbation scheme it will be considered the conformal normal curvature (Poon & Poon 1996) in the direction \( 1 \)

\[
B_1(\theta) = \frac{-1^T \Delta_p(\mathbf{L}_p)^{-1} \Delta_p \mathbf{I}}{\sqrt{\text{tr}(\Delta_p(\mathbf{L}_p)^{-1} \Delta_p)^2}} \theta_1 \omega_0 = \omega_0,
\]

in which \( \mathbf{L}_p = \partial^2 L_p(\theta, \omega)/\partial \theta \partial \omega^T \) is the penalised information matrix and \( \Delta_p = \partial^2 L_p(\theta, \omega)/\partial \theta \partial \omega, \) \( i = 1, \ldots, p \) and \( j = 1, \ldots, q. \) Observe that it is an injective transformation of the normal curvature (Cook 1986) \( C_1 = 2\parallel \Delta_p(\mathbf{L}_p)^{-1} \Delta_p \parallel, \) which is invariant under conformal re-parametrisations and is a standardised measure. Let \( \{e_1, \ldots, e_q\} \) an orthonormal basis of eigenvectors of \( \Delta_p(\mathbf{L}_p)^{-1} \Delta_p, \) with standardised eigenvalues \( \nu_1, \ldots, \nu_q. \) An eigenvector \( e \) is \( q \) influential if \( |B_1| \geq q/\sqrt{n}. \) Let \( \mathbf{E}_j \) be the column vector in \( \mathbb{R}^n \) whose \( \omega \)th entry is 1 and all other entries are 0, which is called basic perturbation vector. In order to analyse the influence of basic perturbation vectors to all influential eigenvectors, consider the arrangement of the absolute values of the standardised eigenvalues by

\[ \nu_{\text{max}} = \nu_1 \geq \ldots \geq \nu_k \geq q/\sqrt{n} > \nu_{k+1} \geq \ldots \geq \nu_n \geq 0 \]

and use \( a_{ij} \) to denote the \( j \)th element of the standardised eigenvector \( \nu_i. \) The aggregate contribution of the \( j \)th basic perturbation vector to all \( q \) influential eigenvectors is \( m(q) = \sqrt{\sum_{i=1}^{n} a_{ij}^2}. \) When \( q = 0 \) it is possible to assess the aggregate contribution of all eigenvectors. In addition, there is a \( q \) sufficiently big to consider the contribution of individual basic perturbation vector to \( e_{\text{max}} \) only. Let \( m(q) \) and \( sm(q) \) be the average mean and standard deviation of the vector \( m(q), \) respectively. A bench-mark to judge largeness may be \( m(q) + c^* sm(q), \) in which \( c^* \) is a selected constant.

In this work, it will be considered the case-weights scheme to assess the local influence. It consist in attribute weights for the observations for the log-likelihood function associated to each group,

\[
L_p(\theta, \omega) = \sum_{i=1}^{n} \omega_i L_i(\theta) - \frac{\lambda}{2} \alpha^T D_i^T D_i \alpha, \tag{5}
\]

in which \( \omega = (\omega_1, \ldots, \omega_n)^T, \) \( 0 \leq \omega_i \leq 1, \) \( i = 1, \ldots, n, \) and the vector of no perturbation is given by \( \omega_0 = (1, \ldots, 1)^T. \)

### 4. Application

The anti-asthmatic substance theophylline was administered to 12 subjects and measured in 11 time points. It is usual for this application to consider the nonlinear function in equation 1, for \( i = 1, \ldots, n \) and \( j = 1, \ldots, m_i \)

\[
\eta_{ij} = d_i \exp(lK_e + lK_a - lC_i) \frac{[\exp(-e^{lK_a x_{ij}}) - \exp(-e^{lK_e} x_{ij})]}{e^{lK_a} - e^{lK_e}},
\]

where the parameters \( lK_a, lK_e \) and \( lC_i \) represent the logarithm of the absorption, elimination and clearance rates and \( d_i \) represents the dose administered to the \( i \)th individual. In this paper, we consider

\[ Z_i = \left[ \frac{\partial \eta(\beta, \omega, x_i)}{\partial lK_e}, \frac{\partial \eta(\beta, \omega, x_i)}{\partial lK_a}, \frac{\partial \eta(\beta, \omega, x_i)}{\partial lC_i} \right] \]
with $\hat{\beta}$ being the least squares estimates of $\beta = (lK_a, lK_c, lC_l)^T$. The smoothing parameter was chosen by Akaike Information Criterion (AIC) method. The best model among the fitted was obtained with $\hat{\lambda} = 0.01$. Figure 1 suggests that individual 1 is a potential influential case.

![Figure 1: Local influence graphical results under case-weight perturbation.](image)

(a) Plot of eigenvalues  (b) Plot of $m(0)$ for $\hat{\theta}$  (c) Plot of $m(2)$ for $\hat{\theta}$

5. Conclusions
We proposed a method for identifying influential observation in data modeled with semiparametric partially nonlinear mixed-effects models. In the presentation, we are going to develop other perturbation schemes such as scale matrix perturbation and response perturbation (Russo et al. 2009).

References


